

Qualitätsgesichertes Datenmanagement für die Sozialforschung

Meier, Friedhelm

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Meier, F. (2003). Qualitätsgesichertes Datenmanagement für die Sozialforschung. *ZA-Information / Zentralarchiv für Empirische Sozialforschung*, 52, 58-71. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-199006>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Qualitätsgesichertes Datenmanagement für die Sozialforschung

von Friedhelm Meier¹

Zusammenfassung

Ein qualitätsgesichertes Datenmanagement auf Abteilungs- oder Institutsebene ist für weite Bereiche wissenschaftlicher Forschung Utopie. Bei dem Bemühen um Qualität und Evaluation, das allorten initiiert wird, ist ein qualitätsgesichertes Datenmanagement allerdings erste Pflicht. Die Aufgaben und Funktionen eines Datenbanksystems, das Qualität und Sicherheit der Daten gewährleistet und zugleich die Studienplanung und Untersuchungsdurchführung im Internet, aber auch konventionell unterstützt, werden am Beispiel des „Aktiven Qualitätsmanagementsystems“ AQS 8i dargestellt.

Abstract

A quality assured data management at departmental or institutional level seems to utopian for wide areas of scientific research. Because quality management and evaluation processes are widely initiated a quality assured data management is a must. The tasks and functions of a database system, which guarantees data quality and data security and in addition supports study development and investigation processes via internet, even conventionally, will be presented as the “Active Quality Management System” AQS 8i.

Datenerfassung und Datenmanagement sind in weiten Bereichen wissenschaftlichen Handelns noch immer individuelle Verantwortungsbereiche, d.h. der Forscher hütet seine Datenmengen auf seinem PC oder in seinem „System Account“. Abteilungen und Institute stellen ihre IT-Ressourcen zur Verfügung, betrachten aber ein qualitätsgesichertes Datenmanagement nicht als ihre Aufgabe bzw. als ihre Zuständigkeit.

¹ Dr. **Friedhelm Meier** ist Professor an der Fakultät für Psychologie der Ruhr-Universität Bochum. E-Mail: Friedhelm.Meier@Ruhr-Uni-Bochum.de

Als Ergebnis dieser allgemeinen Situation finden wir auf den Rechensystemen unendliche Dateimengen, unterschieden nach Namen und Datum, deren Inhalte unterschiedliche Bearbeitungsversionen mit unüberschaubaren Redundanzen darstellen. Hat der Doktorand seine Arbeit abgegeben oder hat der wissenschaftliche Mitarbeiter eine neue Beschäftigung gefunden, dauert es nicht lange, bis die Frage aufkommt: „Wo sind die Daten geblieben?“

1 Funktionen eines qualitätsgesicherten Datenmanagements

Ein qualitätsgesichertes Datenmanagement soll folgende Funktionen sicherstellen:

- Datenerhebung, Datenmanagement, Datenverbreitung werden als zentrale Aufgabe und Verantwortung der Organisation, d.h. der Abteilung, des Institutes, der Körperschaft, unabhängig von Personen verstanden.
- Die Generierung von Daten wird von der Studienplanung bis zur Studienausswertung zentral durchgeführt und dokumentiert.
- Die generierten Datenmengen sind Rohdaten, d.h. Originaldaten, die redundanzfrei und eindeutig zentral gespeichert werden.
- Die Veränderung der Daten nach ihrer Ersteingabe, d.h. nach ihrer originalen Generierung müssen lückenlos dokumentiert werden.
- Veränderte Daten dürfen nicht gelöscht werden, um die Änderungen jederzeit nachvollziehen zu können.
- Der Grund der Änderung und der Verantwortliche müssen mit der Änderung gespeichert werden.
- Die Berechtigung für den Zugriff auf eine Datenmenge und deren Veränderung muss geregelt sein.

Die genannten Aufgaben bzw. Funktionen eines qualitätsgesicherten Datenmanagements ergeben entsprechende technische Anforderungen an das IT-System, das diese Aufgaben erfüllen soll.

Eine wesentliche Anforderung ergibt sich aus der lückenlosen Kontrolle der Generierung und der Bearbeitung eines erhobenen Datums bzw. eines erfassten Messwertes an die Datenhaltung: Auf dieser elementaren Ebene können Daten nicht mehr in Dateien organisiert werden, sondern müssen in Datenbanken gespeichert

werden, die eine Adressierung und flexible Verknüpfung jeden einzelnen Datums² ermöglichen.

In relationalen Datenbanken werden Daten in Tabellen organisiert, die diese elementaren Bezüge ermöglichen.

Es ist nicht besonders sinnvoll, in Datenbanken wieder so viele Tabellen zu erzeugen wie Dateien in der herkömmlichen Dateistruktur. Originaldaten müssen eindeutig, d.h. redundanzfrei, gespeichert werden, um Duplikate und damit unterschiedliche, konkurrierende Änderungszustände zu vermeiden.

Die Datenbank muss zentral implementiert, verwaltet und der Zugriff kontrolliert werden.

Die Datenbank muss zeitlich und räumlich kontinuierlich verfügbar sein, um jede Änderung der Daten und jede Eingabe bzw. Abfrage direkt bearbeiten zu können.

Die Nutzung der Datenbank als zentrale Instanz des Wissensmanagements einer organisatorischen Einheit muss frühzeitig in den Prozess der Datengenerierung, d.h. in den Forschungsprozess, einbezogen werden. Dies bedeutet, dass die in der Datenbank vorhandenen Daten und Wissensmengen wieder in die Studienplanung und Studienkonzeption einfließen und somit strukturierend und normierend wirken.

Die Datenstruktur der Datenbank muss allgemein und strukturäquivalent zu weiteren Datenbanken sein, um Metadatenanalysen einfach und schnell zu ermöglichen.

Dementsprechend müssen alle abgespeicherten Daten über den gesamten Lebenszyklus der Datenbank verfügbar sein und dürfen nicht in sedimentierende Archive ausgegliedert werden. Damit wird zugleich die Möglichkeit geschaffen, alle Daten auf neue Versionen des Datenbanksystems zu migrieren, ohne jeden Datensatz einzeln bearbeiten zu müssen.

Die beschriebenen Funktionen und Aufgaben eines qualitätsgesicherten Datenmanagements können für manchen Forscher bedrohlich wirken, da er sich eingeschränkt fühlen mag. Diese scheinbare Einschränkung bedeutet aber gleichzeitig eine Befreiung von unsystematisch bewältigten Ordnungszwängen.

2 In diesem Zusammenhang muss klargestellt werden, dass hier unter einem Datum immer der empirisch ermittelte bzw. gemessene Wert verstanden wird. Dieser Begriff muss unterschieden werden von dem allgemeinen Informatikbegriff und der kalendarischen Zeitangabe.

2 Das Aktive Qualitätsmanagementsystem AQS

Nachfolgend wird ein speziell für Forschungsprozesse entwickeltes Datenbanksystem vorgestellt, das die gesamten Abläufe und Aufgaben qualitätsgesicherten Datenmanagements realisiert und gewährleistet.

Der Einsatz dieses Instruments beginnt dabei nicht erst mit der vorhin beschriebenen Datenhaltung oder Datensatzpflege. Das System erlaubt vielmehr die Steuerung des gesamten Forschungsprozesses von der Studienkonzeption über die Erhebung bis zur Analyse und Berichtlegung. Es können sogar zeitkontrollierte Erhebungsabläufe definiert und ausgeführt werden. Dazu können Erhebungszeitpunkte vom Minuten- bis zum Tages- und Monatsbereich definiert und ausgeführt werden. So kann der Psychologe minutengenau seine „Recall-Studien“ und der Marktforscher taggenau seine „Follow-up-Befragungen“ realisieren.

Das „Aktive Qualitätsmanagementsystem“ AQS unterstützt sowohl die Forschungsplanung, die zentrale oder dezentrale Datenerfassung und das Datenmanagement als auch aktiv die Steuerung der Untersuchungsabläufe. Es ist also nicht lediglich ein Protokoll- und Dokumentationssystem.

Das AQS basiert auf einem *Allgemeinen Datenmodell*, das es ermöglicht, mit einer begrenzten Zahl von Tabellen alle Anforderungen an die Datenerfassung und Dokumentation eines empirischen Forschungsprojektes redundanzfrei zu modellieren. Diese Tabellen enthalten die Beschreibung der Datenattribute, die Beschreibung der einzelnen Messungen und die empirisch gewonnenen Daten. In ergänzenden Tabellen werden alle Änderungen der Attribute dokumentiert und die Untersuchungsabläufe modelliert.

Obwohl die Anzahl der Datenattribute theoretisch beliebig ist, muss diese konzeptionell doch auf eine sparsame und allgemeingültige Menge beschränkt werden, die es erlaubt, Daten aus den unterschiedlichsten Forschungsbereichen strukturäquivalent abzubilden. Die Attributmenge des hier beschriebenen AQS umfasst die Merkmale Study, Project, Design, Agent, Object, Subject, Sample, Variable, Scale, Scalevalue, Topic, Method und die Benutzermerkmale Operator und User.

Der eine Teil dieser Attribute beschreibt die Untersuchungsfälle:

- Das Attribut *Study* ist das generelle strukturierende Merkmal des gesamten vorhandenen Datenkörpers.
- Das Attribut *Project* hilft, Studieneinträge zu gruppieren und Abfragen von Studien zu vereinfachen.

- Das Attribut *Design* beschreibt die Zugehörigkeit der Daten zu experimentellen oder anderen Untersuchungsbedingungen, etwa „Kontrollgruppe“, „Pre-test“ usw., und kann bei der statistischen Analyse für die Faktorenbildung genutzt werden.
- Das Attribut *Agent* beschreibt das für die Untersuchung zuständige Institut, z.B. in multizentrischen oder interkulturellen Studien, oder den Untersucher oder andere Instanzen, die eine Untersuchung betreiben.
- Das Attribut *Object* beschreibt den Untersuchungsgegenstand oder das Untersuchungsthema, etwa „Reaktionstyp“, „Wertemuster“.
- Das Attribut *Subject* beschreibt die untersuchte Instanz bzw. die Beobachtungs- oder Erhebungseinheit, z.B. den Befragungsteilnehmer, den Freiwilligen in experimentellen Anordnungen.
- Das Attribut *Sample* beschreibt die Zugehörigkeit zu einer Untersuchungsgruppe oder Stichprobe.

Der andere Teil der Attribute beschreibt die Untersuchungsmethoden:

- Das Attribut *Variable* beinhaltet die Frage aus einem Erhebungsbogen.
- Das Attribut *Scale* beschreibt die Maßeinheit, wie „Zustimmung“, „nein/ja“ usw. aber auch „cm“ oder „kg“. Darüber hinaus beschreibt das Attribut das Datenformat des Messwertes, ob es sich um einen Antworttext aus einer offenen Frage handelt oder um einen numerischen Wert, um eine kalendarische Datumsangabe, um eine Uhrzeit oder um eine Instruktion zur Erläuterung der Messung.
- Das Attribut *Scalevalue* beschreibt die Bedeutung einzelner Werte in standardisierten bzw. codierten Antworten.
- Das Attribut *Topic* hilft, Abfragen von Variablen aus großen Variablenmengen inhaltlich zu gruppieren. Große Variablenmengen können die Definitionen mehrerer Tausend Variablen umfassen und können daher zu Inhalts- oder Verfahrensbereichen zusammengefasst werden.
- Das Attribut *Method* gruppiert Variablen zu einem Erhebungs- und Messinstrument. Eine Variable kann dabei in verschiedenen Messinstrumenten verwendet werden. Dadurch wird eine Normierung erzielt, die insbesondere Metaanalysen deutlich präzisiert. Mit der Variablenauswahl können zugleich Wertebereiche und Filtersprünge für den Erhebungsablauf definiert werden. Eine Methode benennt und definiert die Durchführung des Erhebungsinstrumentes.

Für das Datenmanagement sind die übrigen Attribute erforderlich:

- Das Attribut *Operator* kennzeichnet jeden Eintrag in das AQS mit dem verantwortlichen Akteur. Dies betrifft sowohl Ersteingaben als auch Korrekturen. Das Attribut kennzeichnet immer Einzelpersonen bzw. Einzelverantwortliche und stellt eine AQS-Zugriffsberechtigung dar.
- Das Attribut *User* kennzeichnet den Benutzer des Oracle Datenbanksystems und seine Zugriffsberechtigung. Das können lediglich Leserechte oder aber Schreibrechte oder Administratorrechte sein. Das Attribut kennzeichnet nicht ausschließlich Einzelpersonen, sondern kann auch Benutzergruppen meinen.

2.1 Datenerfassung und Datenmanagement mit AQS

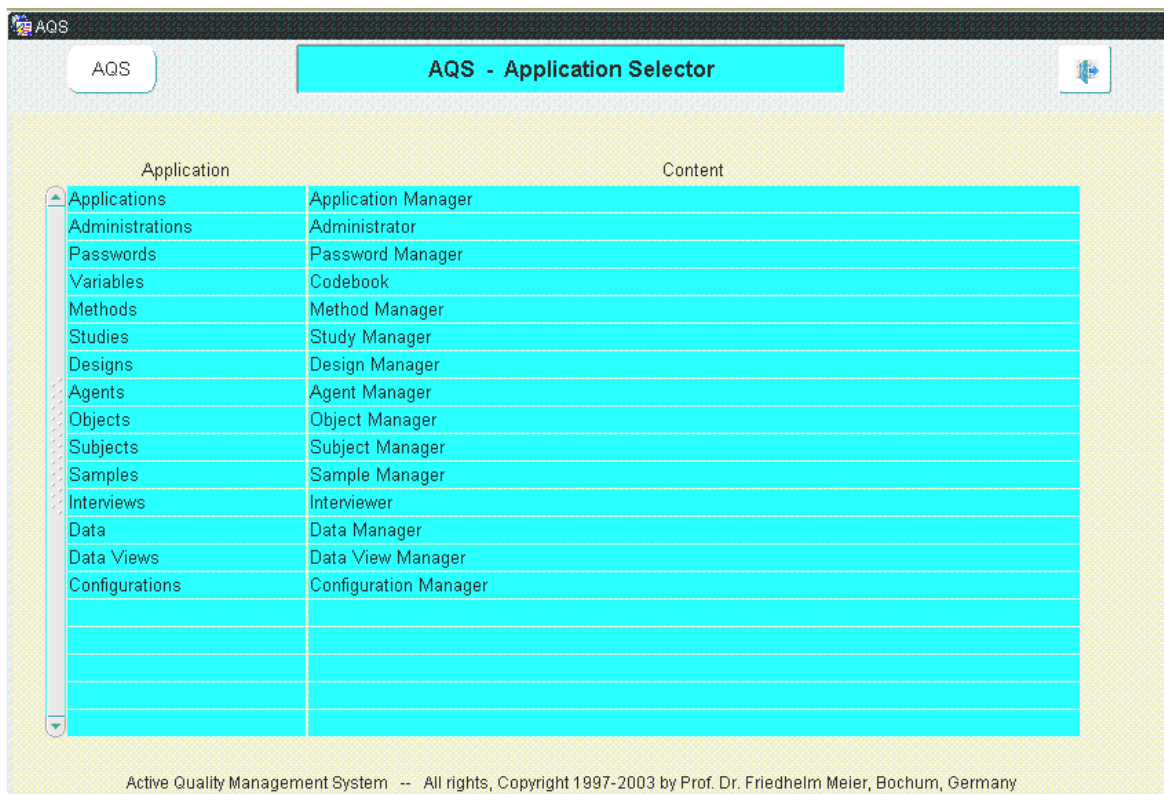
Das AQS wurde kohärent in dem Datenbanksystem Oracle 9i mit der graphischen Benutzeroberfläche Oracle Forms 9i realisiert, um effiziente Werkzeuge für die Entwicklung und Anwendung nutzen zu können. Damit ist auch die Internetintegration bzw. die dezentrale Dateneingabe (*Remote Data Entry*), Datenabfrage und Datenverwaltung systemisch gewährleistet. Zu der umfassenden dezentralen Datenerfassung mittels Internet kommt ebenso die weltweite Abfragemöglichkeit der Daten mittels Berichts- und Analysesoftware, wie statistischer Analyse-, Office- oder Data-mining-Software.

In Abbildung 1 wird die Eingangsmaske des *Application Selector* wiedergegeben, die zum Aufruf der einzelnen AQS-Anwendungen dient, mit denen einerseits die oben beschriebenen Attributdefinitionen durchgeführt werden und andererseits die Datenerhebung mit dem *Interviewer* oder das Datenmanagement mit dem *Data Manager* gestartet werden können. Diese Anwendungen werden nachfolgend kurz beschrieben.

Der *Application Manager*, der *Administrator*, der *Password Manager* und der *Configuration Manager* sind Verwaltungswerkzeuge, welche die Administration des AQS erheblich vereinfachen und den Einsatz der komplexen – und bezüglich der Datenbankkohärenz ungesicherten. – Verwaltungswerkzeuge vermeiden.

Mit den Anwendungen *Study*, *Design*, *Agent*, *Object*, *Subject* und *Sample Manager* sowie *Codebook* und *Method Manager* werden die einzelnen Attribute zur Datenbeschreibung verwaltet.

Der *Data View Manager* erlaubt es, beliebige Variablenmengen zusammenzustellen und für die Datenabfrage und die Berichts- oder Analyseerstellung für Programme, wie MS Office, SAS, SPSS, usw. verfügbar zu machen.

Abbildung 1 Eingangsmaske des Application Selector

Der *Interviewer* ist die zentrale Anwendung für die Erfassung von Originaldaten, die überall im Internet mit unterschiedlichen Funktionen dargestellt und aufgerufen werden kann. Die Anwendung umfasst drei Registerkarten: die erste zur Identifikation des *Operators* oder bei Selbstauskünften des registrierten *Subjects*, die zweite zur Auswahl der Attribute, die auch in der Untersuchungsplanung vordefiniert werden können, die dritte zur Darstellung des Dateneingabeprozesses mit verdeckter Filterführung.

Abbildung 2 zeigt beispielsweise die Registerseite „Input“ des Interviewers. Es wurden auf den vorhergehenden Registerseiten der Operator AQS eingestellt und die Attribute für den „Entry: 1“ ausgewählt (Registerseiten nicht abgebildet). Aktuell wird eine Frage mit 6 standardisierten Antwortreaktionen vorgelegt, die mit der Maus ausgewählt werden können. Der Antwortcode kann aber auch direkt in das Eingabefeld eingetragen werden. Wertebereichsprüfungen werden automatisch durchgeführt.

Die Navigation durch das Interview geschieht mit den beiden Pfeil-Schaltflächen am Anzeigefeld der vorgelegten Fragen. Diese Navigation berücksichtigt auch eine Sprung- bzw. Filterführung, die im *Method Manager* festgelegt wurde.

Abbildung 2 Interviewer, Registerseite „Input“

The screenshot shows the 'AQS - Interviewer' software interface. At the top, there is a navigation bar with buttons for back, forward, and search. Below this, the title 'AQS - Interviewer' and the subtitle 'medizinische Angaben' are displayed. The main area is divided into two tabs: 'AQS' and 'Entry: 1'. The 'Entry: 1' tab is active, showing a question: 'Hatten Sie seit dem letzten Arztbesuch Schmerzen im Brustbereich?'. Below the question, there is a response scale with six options: 'keine', 'schwach', 'deutlich', 'stark', 'sehr stark', and 'unerträglich'. The 'sehr stark' option is selected. At the bottom, there is a table with columns 'Row', 'Response Value', and 'Entry'. The table shows one row with '1' in the 'Row' column, '5' in the 'Response Value' column, and 'sehr stark' in the 'Entry' column.

Row	Response Value	Entry
1	5	sehr stark

Da der *Interviewer* Originaldaten erzeugt, sind dort keine Korrekturprotokolle oder Doppeleingaben zur Kontrolle der Datenübertragung von Papier in das elektronische Medium erforderlich. Anders im nachfolgend dargestellten *Data Manager*: Hier müssen alle Datenkorrekturen begründet bzw. kommentiert werden (vgl. Abbildung 3).

Der *Data Manager* erlaubt eine umfangreiche Abfrage der AQS-Attribute und gibt eine genaue Übersicht über die Eintragsdefinitionen und den Zeitpunkt der Eintragserzeugung und Eintragsbeendigung sowie über die Variablendefinitionen in dieser Methode bzw. in diesem Erhebungsverfahren.

Zusätzlich können mit einer Schaltfläche „invalid“ Daten, das sind zuvor korrigierte oder entfernte Daten, wieder inspiziert werden. Mit einer weiteren Schaltfläche können alle Dateneingaben angezeigt werden, die bereits zu dieser Variablen in der Studie existieren. Ebenso können die Dateneingaben angezeigt werden, die neben der aktuell vorgelegten Variablen bereits in der aktuellen Messung eingetragen wurden. Damit hat der Datenmanager einen umfassenden Überblick über die vorhandene Datenstruktur, die er ergänzen oder korrigieren möchte.

Abbildung 3 Data Manager

The screenshot displays the 'AQS - Data Manager' window. At the top, there are navigation buttons and a title bar. Below the title bar, a tabbed interface shows 'Agent01', 'Design01', 'Object01', 'medfrag2', '99999', and '1'. The 'medfrag2' tab is active, showing a data entry form with fields for 'Response Values', 'Scale', 'Data Type', 'Minimum', 'Maximum', 'Label', and 'Variable'. The 'Label' field contains 'AP01' and the 'Variable' field contains 'Arbeitsplatz01'. Below the form is a table with columns 'Item', 'Content', 'Row', and 'Text'. The table contains three rows of data:

Item	Content	Row	Text
1	Beschreiben Sie wesentliche Aufgaben Ihrer Arbeitstätigkeit.	1	Entwicklung von Datenbankstrukturen
2	Tragen Sie jede einzelne Aufgabe in eine Zeile ein.	2	Analyse von Arbeitsabläufen
3		3	Implementierung von AQS-Systemen

Below the table are buttons for 'Invalid Data', 'Data in Study', and 'Data in Entry'. A 'Reason' dialog box is open in the foreground, showing 'Actual Value' and 'Previous Value' fields. The 'Actual Value' field contains 'Entwicklung von Datenbankstrukturen' and the 'Previous Value' field contains 'Entwicklung von Datenbankkonzepten'. The dialog box also has a 'Give Reason for Correction ...' field and buttons for '<< Back' and 'Reasons'.

Das Protokoll der Änderung von Daten über den Lebenszyklus des Datenbanksystems wird „Audit Trail“ genannt. Es ermöglicht jederzeit, Änderungen nachzuvollziehen und zu prüfen. Eine weitere Funktion des „Audit Trails“ ist die unbegrenzte Möglichkeit, alte Zustände wieder zu aktivieren. Dies wird am Beispiel des „Study Manager“ in Abbildung 4 gezeigt.

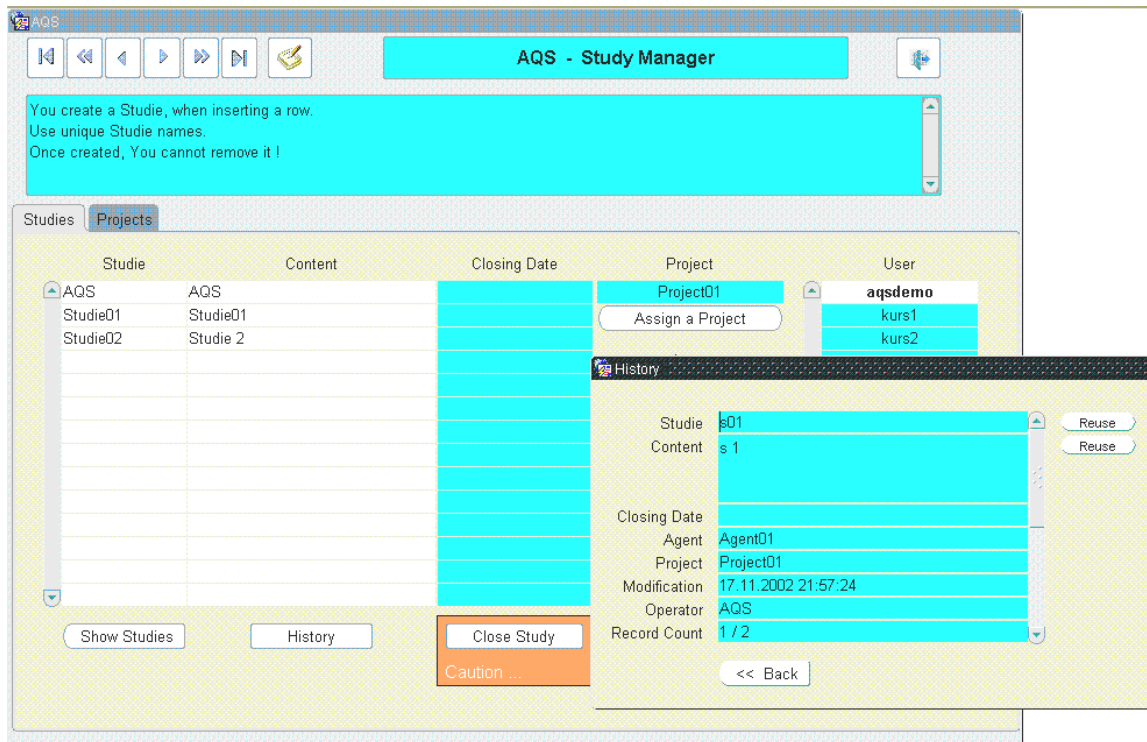
Dort ist die Historie für die Studie 01 geöffnet worden. Es bestehen Unterschiede im Studiennamen und in der Studienbeschreibung (Content). Diese werden durch die sichtbaren Schaltflächen „Reuse“ angezeigt, die zugleich eine Wiederherstellung der Werte erlauben.

Die Abbildung 4 zeigt zudem eine weitere Funktion des qualitätsgesicherten Datenmanagementsystems mit AQS: die Schaltfläche „Close Study“.

Eine Studie muss zu einem festgelegten und dokumentierten Zeitpunkt geschlossen werden. Die Datenerhebungsphase wird mit „Close Study“ abgeschlossen, und somit werden weitere Änderungen verhindert. Dadurch wird der Datensatz für gültige und qualifizierte Analysen und Berichte zugänglich. Wären noch während der Zeit der Berichtlegung Datenkorrekturen möglich, könnten diese Berichte nicht als gültig

genutzt werden. Das AQS verhindert, dass ein geschlossener Datensatz bearbeitet bzw. editiert werden kann. Eine im Wissenschaftsbereich eher unübliche Technik, qualifizierte Analysen und Berichte zu erzeugen.

Abbildung 4 Study Manager



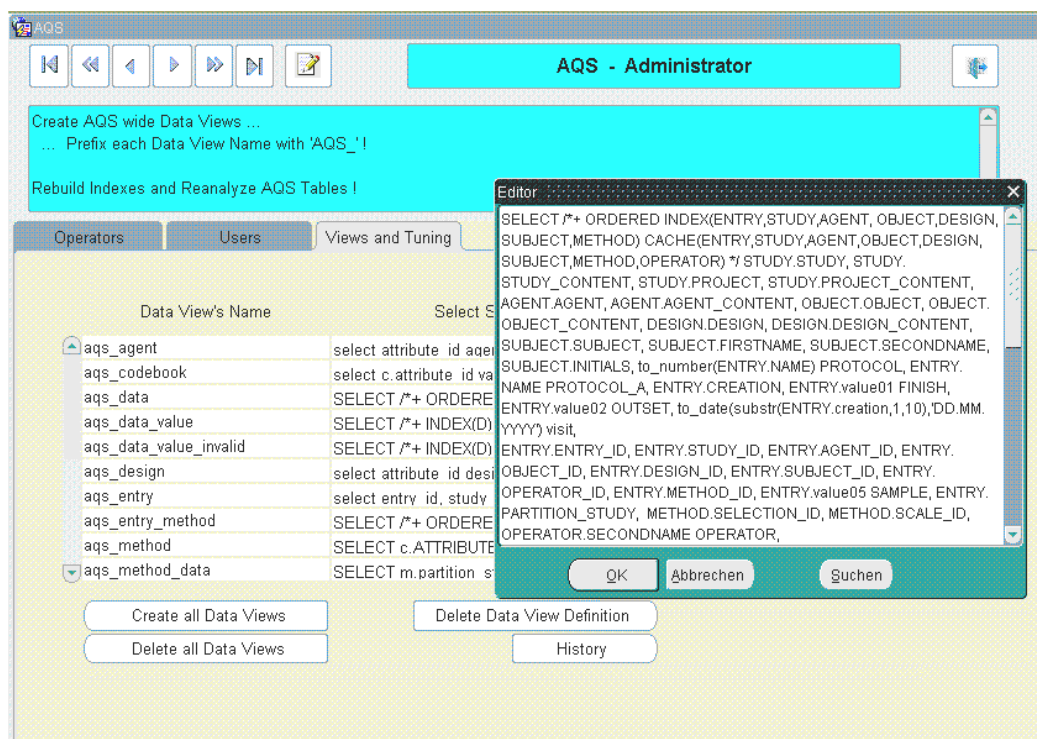
Der Anspruch des AQS, Daten qualifiziert und zentral zu speichern und zu verwalten, erfordert nicht nur, dass die Datenbank von überall erreichbar ist, sondern auch, dass die Anwendungen an die einzelnen Arbeitsplätze angepasst werden können. Dies betrifft insbesondere die Benennung der Felder und Schaltflächen in der Anwendung entsprechend den unterschiedlichen wissenschaftlichen Begriffskulturen. Das AQS selbst ist mit sehr allgemeinen Benennungen ausgestattet, die es jedem Wissenschaftsbereich erlauben, seine Problemstellungen, Studienmerkmale und Variablen zu formulieren.

Dennoch gibt es Vorlieben, denen Rechnung getragen wird. Beispielsweise spricht der eine von „Befragten“, der andere von „Probanden“, der dritte von „Untersuchungsproben“ oder „Verkehrsknoten“. Es ist daher möglich die Benennung „Subject“, wie auch alle anderen Benennungen, für jede Arbeitsumgebung abzuändern. Hierzu können Benennungsprofile vom Administrator erstellt werden, die beim Start einer Anwendung automatisch geladen werden.

2.2 Datenabfrage mit AQS

Ein qualitätsgesichertes Datenbanksystem kann keine Sekundärdaten speichern, da es nur für bereinigte, redundanzfreie Primärdaten, das *Gold* des Instituts, konzipiert ist. Sekundäre Analysedaten, wie Summenwerte, Faktorwerte, Rekodierungen oder gar umstrukturierte Teildatenmengen gehören nicht in einen redundanzfreien Originaldatenkörper.

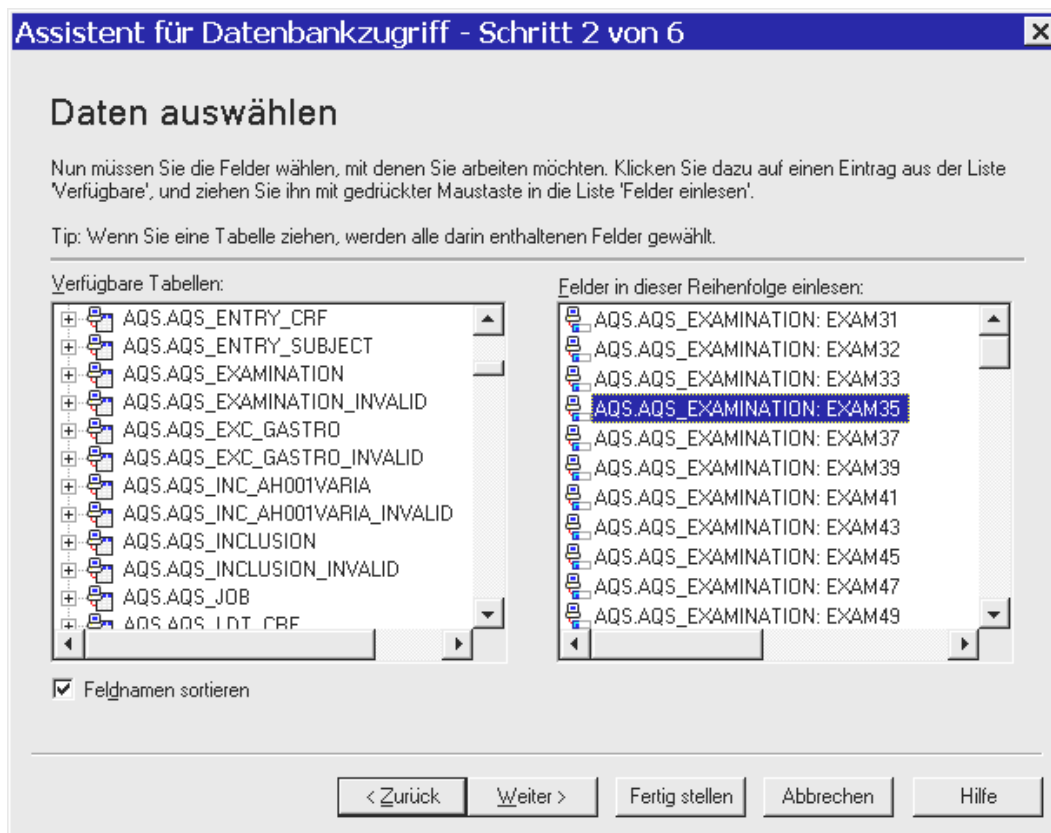
Abbildung 5 Administrator, Registerseite „Views and Tuning“



Viel sinnvoller, als sekundäre Datenmengen zu speichern, ist es, beliebige Abfragealgorithmen zu formulieren und diese in der Datenbank als Datensichten (data views) zu speichern. Die Erzeugungsbedingungen von sekundären Datenkörpern werden dadurch nachvollziehbar dokumentiert und können beliebig genutzt werden.

Das AQS bietet die Möglichkeit, unbegrenzt Datensichten zu definieren und abzuspeichern. Die Datensichten können einerseits im „Data View Manager“ und andererseits im Administrator erstellt werden. Abbildung 5 zeigt die Registerseite „Views and Tuning“ des Administrators mit einer typischen SQL-Anweisung, die mit dem Editor geöffnet wurde.

Die Datensichten werden einerseits in einer Tabelle des AQS abgespeichert und andererseits im Datenbanksystem zur Ausführung erzeugt. Dadurch sind sie über alle Schnittstellen abrufbar, die SQL-Anweisungen transportieren.

Abbildung 6 Assistent für den Datenbankzugriff in SPSS

Microsoft Access, Excel, Word und andere Software, wie etwa SPSS, verfügen über die entsprechenden ODBC-Schnittstellen (Open DataBase Connectivity). Die SAS Software verfügt so gar über eine Oracle-spezifische Schnittstelle, die ein leistungsfähigeres Protokoll verwendet.

Abbildung 6 zeigt den Aufruf von AQS-Datensichten im „Assistenten für den Datenbankzugriff“ im SPSS. Es können hier mehrere Datensichten aufgerufen und im darauf folgenden Schritt zu einer virtuellen Tabelle verknüpft werden. Darüber hinaus können Felder bzw. Variablen umbenannt und Filter für Teildatenmengen definiert werden. Die Daten werden mit dem Fertigstellen des Assistenten wie Daten aus einer Datei ins SPSS geladen und stehen für alle weiteren Bearbeitungsschritte zur Verfügung. Entscheidend ist, dass keine Daten zurückgeschrieben werden können und somit die Kohärenz und Gültigkeit der Originaldaten nicht bedroht sind.

2.3 Der AQS Report Manager

Der AQS Report Manager benutzt auf der Basis des Microsoft Access nur die in AQS bereitgestellten Datensichten und legt selbst keine Datentabellen an.

Der AQS Report Manager liefert Erhebungs- bzw. Messprotokolle. Diese werden entsprechend den Methodendefinitionen automatisch erzeugt und pro Messung ausgegeben. Eine Auswahl der einzelnen Erhebungsprotokolle ist anhand der Attribute des Datenmodells von der gesamten Studie bis zur einzelnen Erhebung bzw. Messung möglich. Die in der AQS Datenbank vorhandenen Variablen- und Methodendefinitionen können jederzeit nach Topics oder Methoden gegliedert abgefragt und ausgegeben werden.

Darüber hinaus ermöglicht es der AQS Report Manager, die üblicherweise internet- oder intranetbasierte Eingabe des AQS mit dem automatisch generierten Ausdruck von Erhebungsbögen für das herkömmliche „Papier-und-Stift“-Protokoll zu ergänzen. Grundlage für die Erstellung sind die eingetragenen und ausgewählten Attribute.

Abbildung 7 Report Manager

The screenshot shows the 'AQS Report Manager - [Record Sheet]' window. At the top, there is a search bar with the text 'Frage hier eingeben'. Below the search bar, the form displays the following information:

- Subject Number:** 99999
- Initials:** A.A
- Period:** 180
- Protocol time [min]:** 180
- Centre:** CRF
- Topic:** Tobacco and Alcohol Use (Social History)

The main section of the form is titled 'Item No. Question' and 'Answer * = mandatory'. It contains several items for data entry:

- Item 1: * Tobacco use**
 - Value:** Mark Your one Answer ...
 - Next Item:** 15
 - Options: 1 Never Used, 2 User, 3 Ex-User, 4 Unknown
- Item 2: * User of Cigarettes**
 - Value:** Mark Your one Answer ...
 - Next Item:** 5
 - Options: 1 no, 2 yes
- Item 3: * Years of Use**
 - Years:** [Text input field]
- Item 4: * Packs per day**
 - f/Day:** [Text input field]
- Item 5: * User of Pipes**
 - Value:** Mark Your one Answer ...
 - Next Item:** [Blank]
 - Options: 1 no, 2 yes
- Item 6: * User of Cigars**
 - [Blank]

At the bottom of the window, there is a 'Seite:' label and a navigation bar with buttons for navigating between pages.

Diese Erhebungsbögen enthalten eine vollständige Filterführung und dienen dann der späteren Dateneingabe mit dem *Data Manager*. Da es sich in dem Fall nicht um ein *Remote Data Entry* bzw. um keine Originaldateneingabe, sondern um die Dateneingabe von einem Protokollblatt bzw. Beleg handelt, muss eine Doppeleingabe durchgeführt werden.

3 Anwendung

Ein qualitätsgesichertes Datenmanagement für die Forschung scheint vor allem bei den neuen Disziplinen des Qualitätsmanagements und der Evaluation von Forschungs- und Lehrbereichen nicht nur wünschenswert, sondern zwingend, um von gesicherten Datensätzen auszugehen, die eine nachhaltige Normierung und damit Metaanalysen über größere Zeitdistanzen ermöglichen. Beispielsweise können Fragebogen und Aufgabenpools im Bereich der Unterrichtsforschung kontinuierlich analysiert und weiterentwickelt werden. In der Sozial- und Marktforschung können Einzelstudien in den gemeinsamen Attributen zu Panelstudien zusammengeführt werden. Dabei stehen alle Daten immer im direkten Zugriff. Das Spiel „Wo ist die Datei?“ kann dann nicht mehr gespielt werden.

Die Erläuterungen zum AQS haben bereits gezeigt, dass der Einsatz eines solchen qualitätsgesicherten Datenmanagementsystems nicht in individueller Verantwortung auf einem Arbeitsrechner, sondern nur auf einem Abteilungs- oder Institutsserver realisiert werden kann. Dieser weist dann eine hohe Verfügbarkeit auf und ist für viele Nutzer weiträumig erreichbar. Dementsprechend sollte die Nutzung sowohl auf der Eingabeseite als auch auf der Abfrageseite hoch genug sein, um eine professionelle Geräteausrüstung zu begründen. So sollten getrennte Server für die AQS-Datenbank, für die AQS-Anwendungen und gegebenenfalls für den Internetzugang installiert werden, um so eine ausreichende Leistung zu bieten. Wenn Abteilungen oder Institute keine ausreichende Auslastung erzeugen, können sie sich wiederum mit anderen zusammenschließen und gegebenenfalls ein zentrales AQS-System in Form des „Data Housing“ nutzen.

Beim „Data Housing“ werden Datenbankdienstleistungen und Rechenkapazitäten vom zentralen Rechenzentrum der Forschungseinrichtung oder auch von einem kommerziellen „Provider“ angeboten.

Der Vorteil eines professionellen zentralen und qualitätsgesicherten Datenmanagements liegt im geringen Administrations- und Ordnungsaufwand an den zahlreichen individuellen Arbeitsplätzen einer Forschungseinheit und der gegebenen benutzerunabhängigen Datensicherung und Datenmigration.

Referenzen

- AQS Active Quality Management System.
Ibis Research Informationstechnologie GmbH & Co.KG, Bochum. www.IBIS-Research.de
- Microsoft Office. Microsoft Corporation, Redmond, WA, USA. www.Microsoft.com
- Oracle 9i. Oracle, Redwood Shores, Ca., USA. www.ORACLE.com
- SAS. SAS Institute Inc., Cary, NC, USA. www.SAS.com
- SPSS. SPSS Inc., Chicago, Ill., USA. www.SPSS.com